Use of role plays to assess therapist competency and its association with client outcomes in psychological interventions: A scoping review and competency research agenda

Katherine E. Ottman[a,1], Brandon A. Kohrt[a,*,1], Gloria A. Pedersen[a], Alison Schafer[b]

[a] Division of Global Mental Health, Department of Psychiatry and Behavioral Sciences, George Washington University, 2120 L St NW, Suite 600, Washington, D.C, 20037, USA
[b] World Health Organization, Department of Mental Health and Substance Use, Avenue Appia 20, Geneva 27, 1211, Switzerland

ARTICLE INFO

ABSTRACT

A major challenge in scaling-up psychological interventions worldwide is how to evaluate competency among new workforces engaged in psychological services. One approach to measuring competency is through standardized role plays. Role plays have the benefits of standardization and reliance on observed behavior rather than written knowledge. However, role plays are also resource intensive and dependent upon inter-rater reliability. We undertook a two-part scoping review to describe how competency is conceptualized in studies evaluating the relationship of competency with client outcomes. We focused on use of role plays including achieving inter-rater reliability and the association with client outcomes. First, we identified 4 reviews encompassing 61 studies evaluating the association of competency with client outcomes. Second, we identified 39 competency evaluation tools, of which 21 were used in comparisons with client outcomes. Inter-rater reliability (intraclass correlation coefficient) was reported for 15 tools and ranged from 0.53 to 0.96 (mean ICC = 0.77). However, we found that none of the outcome comparison studies measured competency with standardized role plays. Instead, studies typically used therapy quality (i.e., session ratings with actual clients) as a proxy for competency. This reveals a gap in the evidence base for competency and its role in predicting client outcomes. We therefore propose a competency research agenda to develop an evidence-base for objective, standardized role plays to measure competency and its association with client outcomes.

Open science registration #: https://osf.io/nqhu7/

1. Background

Studies that form the evidence base for the effectiveness of psychological treatments on client outcomes have historically focused on components of the specific therapies—typically by measuring a provider's adherence or fidelity to distinct treatments. However, the need to ensure that practitioners are providing care completely and effectively has spurred growing interest in training methods and renewed discussion of how to best train providers to deliver psychological treatments. If training methods are to be researched, the outcomes of trainings in psychological interventions (e.g., evaluating whether a provider has the knowledge and skills to deliver the intervention) must be defined and investigated. Determining whether improved provider competency is linked with improved client outcomes is especially important when exploring scalability of interventions worldwide.

Globally, the delivery of psychological treatments by non-specialists (i.e., individuals who lack specialized professional training in fields such as psychiatry, psychology, or clinical social work) in settings within both high-income countries and low- and middle-income countries (LMIC) has been proposed as a solution to reduce the mental health treatment gap (Fairburn & Patel, 2014). This approach of having non-specialists assume some of the roles traditionally performed by mental health professionals is known as task-shifting or task-sharing (Barnett, Lau, & Miranda, 2018; Kakuma et al., 2011; Patel, 2012). A systematic review of psychological treatments delivered by non-specialists for care of persons with common mental disorders identified 25 RCTs, with a standardized mean effect size of 0.48, which is not dissimilar from results achieved by specialists in high resource settings (Singla et al., 2017). Since publication of the review, more recent studies have demonstrated small to large effect sizes for non-specialist

delivered psychological treatments (Bryant et al., 2017; Fuhr et al., 2019; Rahman, Hamdani, Awan, & et al., 2016; Rahman et al., 2019; Sikander et al., 2019). This approach is advocated in the Lancet Commission on Global Mental Health and Sustainable Development as an important innovation for global scale up to reduce the mental health treatment gap (Patel et al., 2018).

Although non-specialists can effectively deliver psychological treatments in both high- and low-resource settings (Montgomery, Kunik, Wilson, Stanley, & Weiss, 2010; van Ginneken et al., 2013), effectively scaling up these services to be widely available for populations in need presents a number of challenges. Two major questions are: what level of skill do non-specialists need to achieve the desired treatment outcomes, and what tools are best suited to evaluate skills to reliably predict treatment outcomes? Therefore, our goal was to explore how competency of therapists (both specialists and non-specialists) has been assessed and compared with client outcomes, with particular attention to use of standardized role plays as one of the advocated competency assessment approaches (Fairburn & Cooper, 2011).

1.1. Therapist competency

There have been ongoing debates about what it means for a therapist to display competency and how it should be assessed (Sharpless & Barber, 2009). (Fairburn & Cooper (2011)) define therapist competency as "the extent to which a therapist has the knowledge and skill required to deliver a treatment to the standard needed for it to achieve its expected effects". Competency, therefore, may be assessed through observable skills of providers working in controlled conditions (e.g., training or supervision sessions) with the use of standardized role plays. Other concepts associated (and often conflated) with therapist competency include therapy quality, treatment adherence, and treatment fidelity. For the purpose of this review, we have included a table of working definitions (Table 1) for clarity of terminology.

Therapy quality is distinguished from therapist competency and described as "the extent to which a psychological treatment was delivered well enough for it to achieve its expected effects" (Fairburn & Cooper, 2011). Therefore, therapist *competency* relates to the knowledge and skills acquired by a provider, while therapy *quality* refers to the way the treatment is delivered. Therapy *quality* may be measured as an observable performance of therapists working in *real-world* conditions (Muse & McManus, 2013). Manualized psychological treatments have traditionally focused on the degree to which therapists are delivering the theory-driven techniques or methods of an intervention (i.e., therapist *adherence*). This is often used interchangeably with the terms "treatment integrity" and "treatment fidelity" (Webb, Derubeis, & Barber, 2010).

These concepts can be mapped onto Miller's original hierarchy of clinical skills proposed three decades ago, which includes four levels (Miller, 1990). In Miller's hierarchy, "knows" refers to a therapist's knowledge as assessed through multiple-choice questions, while "knows how" refers to their ability to apply knowledge, which could be assessed through open-ended responses and/or written clinical vignettes. "Shows" and "does" refer to therapist *competency* and *quality*, respectively.

At the heart of these different concepts is the question: What matters to achieve desired client outcomes? Despite decades of psychological treatment research, there is still an empirical question of whether or not competency matters. Intuitively, therapists with greater competency should have better treatment outcomes than therapists with limited competency (Barber, Sharpless, Klostermann, & McCarthy, 2007). Improving Access to Psychological Therapies (IAPT) National Collaborating Centre for Mental Health, 2019, the Global Social Service Workforce Alliance (GGSWA) (2017), University College London's Centre for Outcomes Research and Effectiveness (CORE) (2019), and others have also proposed detailed competency frameworks built on the expectation that promoting these competencies will lead to quality care and positive client outcomes.

Therefore, as the field of global mental health moves toward expanding availability of psychological treatments in low-resource settings around the world (Fairburn & Patel, 2014), it is an important moment to examine the literature on competency to determine the evidence base for the relationship between competency and outcomes among clients. We conducted a scoping review to document how competency has been operationalized, the types of tools used to measure competency, the types of study designs used to compare competency with treatment outcomes, and the findings to date related to competency and outcomes. The results of this review will be used to inform ongoing efforts in the field of global mental health to improve research and public health initiatives related to scaling up psychological treatments around the world. This review is part of the World Health Organization (WHO) initiative Ensuring Quality in Psychological Services (EQUIP) to develop an open-access platform for resources and tools to conduct competency-based trainings in mental health and psychosocial support (Kohrt et al., 2020).

2. Methods

We chose to conduct a scoping review given the diverse applications that this approach can serve in mapping a field of literature (Levac, Colquhoun, & O'Brien, 2010). Scoping reviews "aim to map rapidly the key concepts underpinning a research area and the main sources and types of evidence available, and can be undertaken as stand-alone projects in their own right, especially where an area is complex or has not been reviewed comprehensively before" (Mays, 2001). There are

Table 1
Working Definitions related to Competency and Client Outcomes.

Term	Definition
Adherence	The degree to which a provider is delivering the theory-driven techniques or methods of a treatment (Webb et al., 2010).
Competency	The observable skills of providers working in controlled conditions; e.g., training or supervision sessions evaluated through the use of role plays (Fairburn & Cooper, 2011).
Quality	The extent to which a psychological treatment was delivered well enough for it to achieve its expected effects. This observable performance of providers working in real-world conditions is evaluated with methods such as clinical observation, clinical notes, and recordings (Fairburn & Cooper, 2011).
Fidelity	The methodological strategies used to monitor and enhance the reliability and validity of behavioral interventions (Resnick et al., 2005).
Therapeutic Alliance	The relationship between therapist and client that consists of three components: the bond, the agreement about the goals of therapy, and the agreement about the tasks of therapy (Wampold, 2015).
Common Factors	These are the factors that all therapies have in common, such as the alliance between the patient and the therapists, creation of expectations (i.e., a rationale that helps patients understand why they have problems and what can be done about them), and enactment of health promoting actions (Cuijpers, Reijnders, & Huibers, 2019; Wampold, 2015). Common factors are interrelated and they often overlap with specific practice elements (Mulder, Murray, & Rucklidge, 2017). Other terms for similar concepts include "nonspecific" or "universal" factors, "core competencies", or "basic" or "foundational" helping skills.
Treatment-Specific Factors	Procedural techniques unique for a particular treatment, often detailed in treatment manuals (Mulder et al., 2017).

four potential reasons to conduct scoping reviews: 1) to examine the extent, range, and nature of research activity; 2) to determine the value of undertaking a full systematic review; 3) to summarize and disseminate research findings; and 4) to identify research gaps in the existing literature (Arksey & O'Malley, 2005). For the purposes of this review, our objectives aligned with reasons 1, 3, and 4.

We followed the Arksey and O'Malley (2005) 6-stage framework for conducting scoping reviews: Stage 1: Identifying the research question (s); Stage 2: Identifying relevant studies; Stage 3: Study selection; Stage 4: Charting the data; Stage 5: Collating, summarizing and reporting the results; and Stage 6: Consulting with relevant stakeholders. Our protocol was drafted using the Preferred Reporting Items for Systematic Reviews and Meta-analysis Protocols: Extension for Scoping Reviews (PRISMA-ScR) (Tricco et al., 2018). The final protocol was registered with the Open Science Framework on 14 February 2019 (https://osf.io/nqhu7/).

2.1. Stage 1: Identifying the research questions

Our scoping review consists of two objectives:

1) An investigation of existing reviews that explore therapist competency in relation to client outcomes; and
2) A review of existing therapist competency assessment tools in the literature that were used to evaluate therapist competency and report client outcomes.

2.2. Stage 2: Identifying relevant studies (search strategy)

For Objective 1 (on the relationship between therapist competency and client outcomes), MEDLINE and Google Scholar were initially searched on July 15, 2018 using the sample search strategy below (Table 2). Objective 1 was structured as a review of reviews. These searches were updated repeatedly with the most recent search completed on February 9, 2019. For the evidence base for the collected assessment tools generated from the above search, the tools' references and citation sources were further reviewed for studies that utilized the tools and evaluated subsequent client outcomes. Final search results were exported into Mendeley, and duplicates were removed. The electronic database searches were supplemented by scanning existing paper references for further relevant studies. Titles and abstracts were then screened for inclusion and exclusion criteria, followed by full-text reviews of the remaining studies to confirm that the studies were relevant and met all criteria. A final, in-depth review was conducted with all articles gathered (see Online Supplemental Table S1 for a list of search terms and records identified).

For Objective 2 (identifying existing competency assessment tools and their relationships with client outcomes), we predetermined the classes for psychological treatment in the search terms. These were "Cognitive Behavioral Therapy (CBT)," "Interpersonal Therapy (IPT)," "Motivational Enhancement," "Behavioral Activation," "Emotional," "Problem-Solving," "Trauma," and "Psychological First Aid (PFA)". We searched Medline, EBSCO, and PsycINFO databases first on July 18, 2018. This search was repeated with the most recent update on February 9, 2019. Sample search strings included "(competen*) OR (assessment*) OR (tool*) OR (rating scale)", along with specific terms for the treatment classes, such as "(Motivational enhancement) OR (Motivational Interview*)". The search was repeated for each psychological class (see Online Supplemental Table S2 for a sample tool search).

Google Scholar was utilized for identifying additional relevant articles, using themes of "therapist competency," "rating scale," "fidelity," and "therapist evaluation" along with the above class categories. The reference lists of all included papers were searched to identify any further manuals or relevant articles, which were subjected to the same screening and selection process.

Table 2
Competency assessment tool characteristics.

Tool Characteristics	Descriptives	Number of Tools (n%)
Number of Items	≤10	11 (28%)
	11 to 20	11 (28%)
	21 to 30	9 (23%)
	30 +	8 (21%)
Provider Type	Non-Specialists	6 (15%)
	Specialists (psychologists/psychiatrists)	33 (85%)
Target Condition	Depression	14 (36%)
	Substance Abuse	15 (38%)
	Other	10 (26%)
Target Population	Adults	30 (77%)
	Children/Adolescents	7 (18%)
	Families/Caregivers	2 (5%)
Likert Response Options	≤3	1 (2%)
	4	3 (8%)
	5	12 (31%)
	6	3 (8%)
	≥7	20 (51%)
Anchoring used for Rating Systems	All anchored (description for every level)	21 (54%)
	Partial anchoring (e.g., extremes and/or midlevel)	10 (26%)
	Minimal anchorings (e.g, only descriptions of "high" or "low")	8 (20%)
Format of Evaluation	Structured role play with standardized client	8 (20%)
	Observed or recorded clinical session	34 (87%)
	Other	0
Comparison with Client Outcomes	Client outcomes compared with structured role plays	0
	Client outcomes compared with observed or recorded clinical sessions	17 (44%)
	Other comparison approach	0
	No comparison with client outcomes	22 (56%)

2.3. Stage 3: Study selection (inclusion and exclusion criteria)

For Objective 1, papers focusing specifically on how assessments of therapist competency were directly measured against the respective client outcomes reported in each study were included. Peer-reviewed journal papers were included if they were published between 1980 and 2019, were written in English, provided a measurement of client mental health outcomes (e.g., measurement of major depressive symptoms 3-months post-treatment), and conducted an assessment of therapist competency using a structured tool. There was no restriction by world region; all low-, medium-, and high-income country-based studies were eligible for inclusion. For Objective 1 (therapist competency and client outcomes), we limited inclusion to studies classified as reviews.

For Objective 2, the review of competency assessment tools, we included all studies that mention a standardized assessment or tool for evaluating therapist competency in their procedures. Assessment tool manuals and/or studies that specified their evaluation tool items (but did not necessarily include a manual) were included. Instruments were included in our item generation procedure if they addressed a psychological therapy within the realm of the above listed treatment classes. Exclusion criteria were studies that did not explicitly detail how therapist competency was evaluated, tools that were solely client-based assessments of therapy, and tools that were solely based on therapist *alliance* or *adherence* to treatment guidelines.

Studies included in the meta-analyses were double-coded with a second reviewer with an agreement above 0.8.

2.4. Stage 4: Charting the data

In accordance with the fourth stage of scoping reviews (Arksey & O'Malley, 2005), selected articles were charted in Microsoft Excel 2016, using a template developed by the team. For Objective 1, the studies included within the reviews were collated into a chart to map out the included studies and to detect duplicates (Supplemental Table S3). For the review of competency assessment tools (Supplemental Table S4), the 39 collected tools were charted with the headings "Assessment Tool," "Author(s)", and "Date." Tool characteristics (such as number of items, provider type, target conditions and populations, as well as anchoring for rating systems) were summarized for comparison and charted (Table 2). Finally, the evidence of outcomes in relation to individual competency assessment tools was charted (Supplemental Table S5) using the following headings: "Assessment Tool; " "Study" where the assessment tool was used to evaluate therapist competency/quality in relationship to outcomes; "Setting," indicating whether the study was conducted in a high-income country (HIC) or a low-to middle-income country (LMIC); "Type of provider," indicating whether the therapist(s) involved in the study were professionals or non-specialists; "Population," indicating the study's stated target population(s); "# of Items" that are on the assessment tool checklist; "Likert scale," or the number of available responses on each checklist item; "Competency Assessment" for indicating how (and if) the study evaluated therapist competency; "Assessment Sample Collection," describing the methods for how the study gathered overall data on the competency/quality evaluations; "Quality Assessment" for indicating how (and if) the study evaluated therapist quality; "Primary Client Outcome(s)" for identifying the main client outcomes investigated by the study; and "Connection to client outcomes," which details the study's main results.

2.5. Stage 5: Collating, summarizing and reporting the results

The fifth stage of a scoping review is to summarize and report the findings (Arksey & O'Malley, 2005). For the review of reviews, the statistical analyses linking therapist competency to client outcomes are reported and discussed. The evidence gathered on individual studies' investigations of therapist competency and client outcomes are further collated, compared, and summarized, and the current state of the literature and recommendations for the field are presented.

3. Results

For Objective 1 (reviews of therapist competency and client outcomes), 88 citations were identified from searches of electronic databases and hand searching article references for existing literature reviews (Fig. 1). Based on the title and the abstract, 75 were excluded with 13 full-text articles retrieved and assessed for eligibility. Of these, a further 9 were excluded due to articles 1) not including more than one study for analysis, 2) focusing on the analysis of therapist *alliance* rather than *competency*, and 3) providing commentary on variations of therapist competency evaluation tools without mention of how the use of these tools (which deemed a therapist "competent") was associated with client outcomes. The remaining 4 studies were considered eligible for this review (Barber et al., 2007; Collyer, Eisler, & Woolgar, 2019; Kazantzis, 2003; Webb et al., 2010).

For Objective 2 (competency assessment tools), 1,932 papers' titles and abstracts were reviewed before the final 117 full-text articles were assessed for eligibility (Fig. 2). Several manuals that were mentioned in the full-text articles were obtained from authors of original studies. Thirty-nine assessment tools (in the form of manuals or papers that detailed the assessment tool items in the absence of an official manual) were included in the final tool review.

To investigate whether the gathered competency assessment tools were used to evaluate client outcomes, studies that cited the 39 assessment tools were searched and yielded 935 citations. Based on title

and abstract review, 68 were retrieved and assessed for eligibility through full-text review. During full-text review, 47 full-text articles were excluded due to the studies not having 1) the documented use of an assessment tool to evaluate therapist competency, and 2) not directly measuring client outcomes. The remaining 21 studies were included for this review.

3.1. Objective 1. Review of Reviews for competency and client outcomes

The scoping review yielded four existing literature reviews, encompassing a combined total of 61 studies that were solely from HIC settings (See Supplemental Table S3). Many studies gathered in the reviews were used in multiple articles and duplicated in our searches for individual competency assessment tools for Objective 2 (described below). While two of the reviews were strictly narrative syntheses that primarily addressed interventions for adults (Barber et al., 2007; Kazantzis, 2003), the remaining two were meta-analyses that investigated elements of therapist competency in psychological interventions targeted for adults (Webb et al., 2010) and children and/or adolescents (Collyer et al., 2019). Of the two meta-analyses, Webb et al. reviewed studies that primarily consisted of cognitive behavioral therapies and dynamic therapies for adults with Major Depressive Disorder (MDD), survivors of child abuse, bulimia nervosa, substance abuse, and a variety of other diagnoses (Webb et al., 2010). Collyer et al.'s studies were primarily cognitive behavioral therapies and parental/familial behavioral interventions targeting children and adolescents with substance abuse, depressive symptoms, anxiety disorders, or behavioral problems (Collyer et al., 2019).

3.1.1. Operationalization of competency

All four of the reviews varied in the operationalization of therapist competency, adherence, and fidelity, and subsequent comparisons of client outcomes. Barber et al. (2007) suggested that "competency" could be divided into two separate categories: "Global competency," or the idea that a therapist has a broad range of abilities to manage problems and assist patients with realizing their goals, closely aligns with our concept of *common factors*; the second category, "limited-domain competence," is a subset of global competency that refers to specific practices for particular interventions, e.g., treatment-specific factors (Barber et al., 2007). All reviews noted that while treatment *adherence* is a vital component of such "limited-domain competence," its measurement has often been the standalone indicator of therapist competency in intervention studies. As such, the two meta-analytic reviews by Webb et al. and Collyer et al. included separate analyses for adherence-outcome studies and competency-outcome studies to determine their respective effects on client outcomes (Collyer et al., 2019; Webb et al., 2010). The most recent review by Collyer et al. (2019) went further and investigated "composite treatment fidelity" as a predictor of client outcomes, defining it as a composite of both treatment adherence and competence for the purpose of their study ($k = 9$).

The two narrative syntheses primarily discussed the operationalization of therapist competency and adherence, examining the psychometric properties of several different assessment tools and describing the outcomes of studies that used these tools for evaluation of therapist competency. Each review concluded that they found overall positive, albeit weak, associations between therapist competency and client outcomes (Barber et al., 2007; Kazantzis, 2003).

None of the included studies used competency assessment as recommended by (Fairburn & Cooper (2011)), i.e., knowledge tests or standardized role plays. Instead, therapy quality, i.e., rating by an observer or supervisor of actual sessions with clients, was used as a proxy for competency. This was primarily evaluated through ratings of recorded sessions by third party raters (please see Supplemental Table S5 for further details).
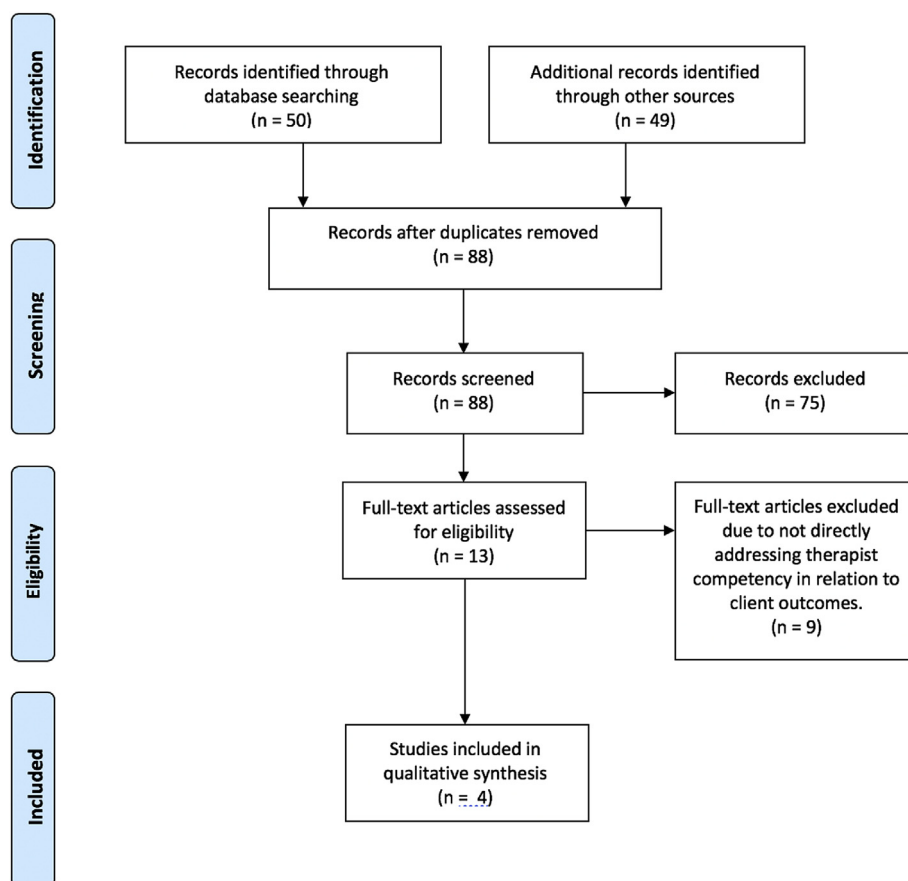
Fig. 1. Search strategy, therapist competency reviews.

3.1.2. Summary of meta-analysis findings

In their meta-analysis, Webb et al. (2010) used correlations between adherence or competency measures and clinical outcomes to calculate the estimated mean population effect and its significance as an average of standardized Fisher's z transformed effect sizes, weighted by the inverse of their variance. They divided their included studies into those that investigated therapist adherence in relation to client outcomes and those that investigated therapist competency. Positive effect sizes would indicate that higher adherence or competency ratings were associated with better outcomes. With a mean weighted adherence-outcome effect size of $r = 0.2$ ($z = 0.36$, 95% CI [-0.069, 0.100], $p = 0.72$, $k = 30$) and a mean weighted competency-outcome effect size of $r = 0.07$ ($z = 0.97$, 95% CI [-0.069, 0.201] $p = 0.33$, $k = 16$), the authors concluded that variability in neither adherence nor competency was significantly associated with client outcomes (Webb et al., 2010).

In a more recent systematic review of psychological interventions targeted for children and adolescents, Collyer et al. (2019) also used correlations for their analyses. They categorized their studies between those investigating therapist adherence or competency as well, but they added a third "fidelity" category for those studies that were reporting a composite of these two constructs. Similarly to the meta-analysis conducted by Webb et al. (2010) they extracted correlation values and conducted z tests. The average weighted adherence-outcome effect size indicated a small but statistically significant relationship between therapist adherence and client outcomes, $r = 0.096$ ($z = 4.938$, 95% CI [0.058, 0.134], $p < 0.001$, $k = 29$). For the competence-outcome effect sizes, competence did not have a statistically significant association with client outcomes, $r = 0.026$ ($z = 1.119$, 95% CI [-0.020, 0.073], $p = 0.263$, $k = 9$). Fidelity (i.e., their composite competency-adherence category) also did not have a significant association with client outcomes, $r = 0.06$ ($z = 0.9153$, 95% CI [-0.070, 0.191],

$p = 0.360$, $k = 5$).

Both meta-analyses conducted homogeneity analyses based on Hedges and Olkin's Q statistic (Hedges & Olkin, 1985) to test whether the observed variability across effect sizes is greater than expected from subject-level sampling error. In their analysis, Webb et al. (2010) reported that the effect size distributions for both adherence to client outcomes ($Q = 50.90$, $p < 0.01$, $k = 30$) and competence to client outcomes ($Q = 37.15$, $p < 0.01$, $k = 16$) were both significantly heterogeneous, suggesting that their variability is greater than expected from sampling error alone. To further examine whether the heterogeneity of effect sizes between studies were due to between-study variability rather than sampling error alone, the authors additionally conducted an I^2 test (Higgins, Thompson, Deeks, & Altman, 2003). A moderate to high I^2 value was associated with both the adherence-outcome (47%) and the competence-outcome (59.6%) effect sizes, indicating that approximately half of the variability in effect sizes across both the adherence and competence studies was due to differences relating to aspects of these studies, rather than sampling error alone.

They further conducted moderator analyses for treatment modality, targeted mental health conditions, and whether temporal confounds (i.e., whether the symptom change that preceded the assessment of competency had been accounted for, methodologically or statistically) or therapeutic alliance confounds were controlled in the studies. There were no moderation effects observed for adherence to client -outcomes' effect sizes, but mean weighted effect sizes for competence-outcome relations were significantly different across the types of problems targeted ($Q = 19.33$, $p < 0.001$) and significantly smaller when the therapeutic alliance was statistically controlled ($rs = -0.03$ and 0.00) compared to when it was not ($rs = 0.23$ and 0.26, $Q = 4.58$, $p < 0.05$). These findings suggest that the types of problems presented by clients were more strongly related to client outcomes than the
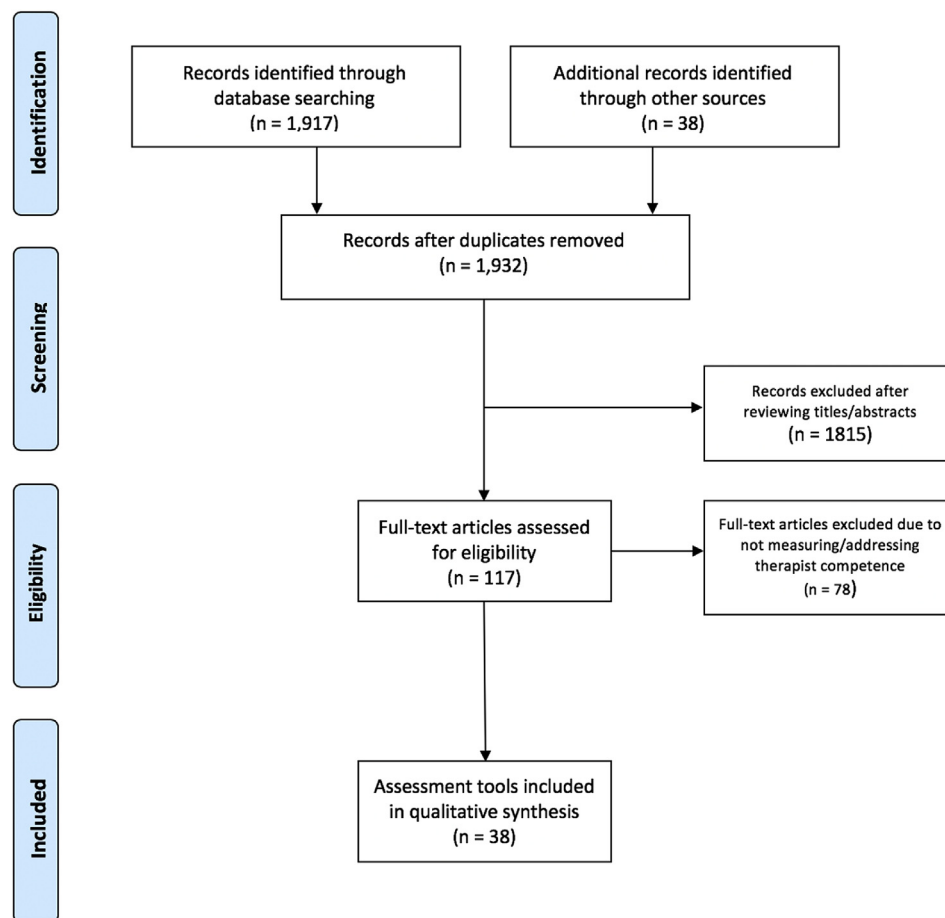
Fig. 2. Search strategy, competency assessment tools review.

therapeutic alliance between provider and client.

For their analyses of adherence-outcome effect studies, Collyer et al. (2019) reported a variance in effect sizes that was significantly greater than would be expected by sampling error alone ($Q = 62.352$, $p < 0.001, k = 43$). There was no significant variance in effect sizes of competence to client outcome relations ($Q = 2.595, p = 0.957$), nor fidelity to client outcome ($Q = 7.700, p = 0.103$) respectively, indicating the studies likely represent a common population mean. They additionally conducted moderator analyses for adherence-outcome and competence-outcome effect sizes (analysis on the fidelity-outcome studies was not feasible due to small sample size). There were no moderation effects observed based on clinical groups, intervention types, and implementation measure informants.

3.2. Objective 2. Review of competency assessment tools and client outcomes

Our search for manuals and studies with competency assessment tools generated a total of 39 tools used in 43 studies (See **Online Supplemental Tables S4 and S5**). The vast majority were developed for use with evaluating specialist providers (85%, $n = 33$), see Table 2. Most tools were designed for use with adult clients (77%, $n = 30$) and the targeted conditions for treatment were primarily substance abuse (38%, $n = 15$) and depression (36%, $n = 14$). The number of items included in each tool varied widely from 1 item up to 99 assessment items (Fig. 3), with the predominant number of tools having 20 items or fewer (56%, $n = 22$). The majority of tools had 7 or more Likert response options (51%, $n = 20$), with 5 response options used in one third of the tools (31%, $n = 12$). Interestingly, while most tools had rating systems that were completely anchored (54%, $n = 21$), there

were still relatively large numbers of tools that only had partial anchoring for their Likert responses (26%, $n = 10$) or minimal descriptions (20%, $n = 8$). Inter-rater reliability for the assessment tools was reported for 15 of the instruments. The inter-rater reliability (intra-class correlation coefficient) ranged from 0.53 to 0.96 (mean, 0.77).

3.2.1. Operationalization of competency

Among the 43 studies, only 8 used a structured role play consistent with (Fairburn & Cooper (2011)) recommended evaluation of competency. None of these 8 studies compared these to client outcomes. All studies that did involve a comparison (see description below) with client outcomes used therapy quality (i.e., rating of actual sessions) as a proxy for competency.

3.2.2. Association of competency assessment tools and client outcomes

Seventeen of the competency assessment tools were used in twenty-one studies for evaluating associations with client outcomes. After gathering these tools, we searched for studies that utilized them for assessing therapist competency in conjunction with evaluating client outcomes. Of the 21 studies forming the evidence base around our collected competency assessment tools, 17 were conducted in HICs by specialist therapists, 1 was conducted in a HIC with a non-specialist counselor (Butler et al., 2013), and 3 were conducted in LMIC settings with non-specialist counselors (Chowdhary et al., 2016; Nadkarni et al., 2017; Papas et al., 2011).

Among all studies from HIC settings, the evidence is mixed; 8 studies indicated a positive association between therapy quality and client outcomes (Campos-Melady, Smith, Meyers, Godley, & Godley, 2017; Forgatch, Patterson, & DeGarmo, 2005; Gaume, Gmel, Faouzi, & Daeppen, 2009; Ginzburg et al., 2012; Hoffart, Sexton, Nordahl, &
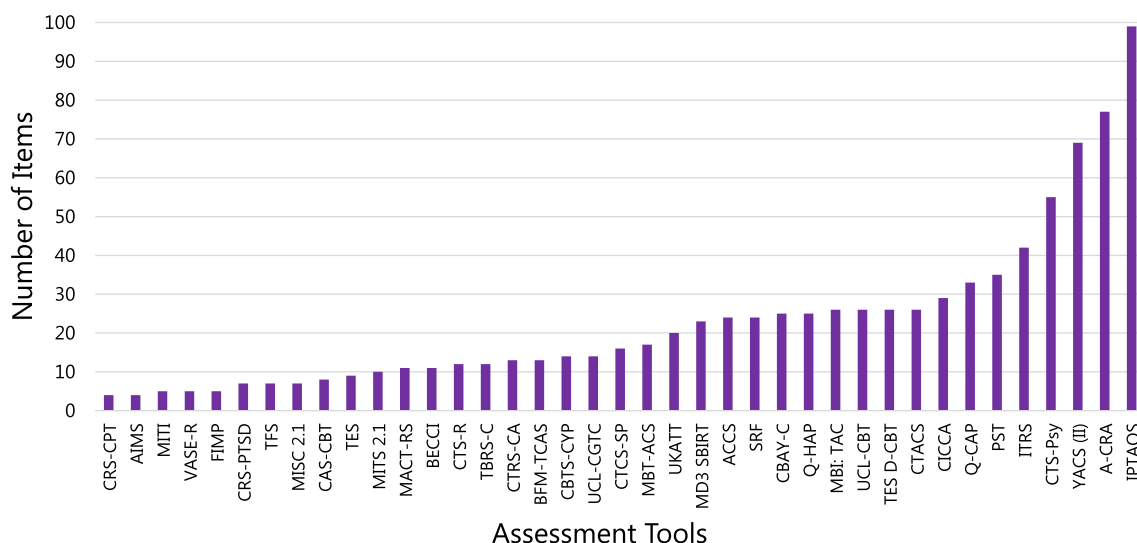
Fig. 3. Competency Assessment Tool Items.

Abbreviations: Alcohol Intervention Mechanisms Scale (AIMS); Assessment of Core CBT Skills (ACCS); Behavioral Family Management Therapist Competency/Adherence Scale (BFM-TCAS); Behavior Change Counselling Index (BECCI); CICCA Scale; Cognitive Therapy Adherence-Competence Scale (CTACS); Cognitive Therapy Competence Scale for Social Phobia (CTCS-SP); Cognitive Therapy Rating Scale for Children and Adolescents (CTRS-CA); Cognitive Therapy Scale for Psychoses (CTS-Psy); Cognitive Therapy Scale Revised (CTS-R); Cognitive-Behavioral Treatment for Anxiety in Youth Competence Scale (CBAY-C); Competence and Adherence Scale for Cognitive Behavioral Therapy (CAS-CBT) for anxiety disorders in youth; Competence Rating Scale for Cognitive Processing Therapy (CRS-CPT); Competence Rating Scale for PTSD (CRS-PTSD); Fidelity of Implementation Rating System (FIMP); Independent Tape Rater Scale (ITRS); IPT Adherence and Quality Scale (IPTAQS); Manual -Assisted Cognitive Behavior Therapy Rating Scale (MACT-RS); Manual for the Motivational Interviewing Skill Code Version 2.1 (MISC); Mentalization-Based Treatment for BPD (MBT-ACS); Mindfulness-based Interventions Teaching Assessment Criteria (MBI:TAC); Motivational Interviewing Target Scheme (MITS 2.1); Motivational Interviewing Treatment Integrity 3.1.1 (MITI 3.1.1); Guided Treatment Software to Teach Clinicians How to Deliver Problem-Solving Treatment for Depression (ePST); Quality of the Counselling for Alcohol Programme (Q-CAP) scale; Quality of the Healthy Activity Programme (Q-HAP) scale; Screening, Brief Intervention, and Referral to Treatment Coding Scale (MD3 SBIRT); Supervisor Rating Form (SRF); The Cognitive Behavior Therapy Scale for Children and Young People (CBTS-CYP); Therapeutic Competence Scale (TCS) for Developmentally Adapted Cognitive Processing Therapy for children and adolescents (D-CPT); Therapist Behavior Rating Scale–Competence (TBRS-C); Therapist Empathy Scale (TES); Therapist Facilitating Scale (TFS); UCL Scale for Assessing Cognitive and Behavioral Therapy (UCL-CBT); UCL Scale for Assessing Core and Generic Therapeutic Competences (UCL-CGTC); UK Alcohol Treatment Trial Process Rating Scale (UKATT); Video Assessment of Simulated Encounters (VASE-R); Yale Adherence and Competence Scale Guidelines (II) (YACS).

Stiles, 2005; Martino, Ball, Nich, Frankforter, & Carroll, 2008; Möller, Karlgren, Sandell, Falkenström, & Philips, 2017; Trepka, Rees, Shapiro, Hardy, & Barkham, 2004), 5 reported mixed results (Carlson & de Mamani, 2010; Davidson et al., 2004; Gutermann et al., 2015; Paivio, Holowaty, & Hall, 2004; Strunk, Brotman, DeRubeis, & Hollon, 2010), and 5 reported no effect (Butler et al., 2013; Gaume, Gmel, & Daeppen, 2008; Hogue et al., 2008a; Shaw et al., 1999; Spohr, Taxman, Rodriguez, & Walters, 2016). Of those with mixed results, most of the studies reported a positive association with some primary outcomes (i.e., measured treatment targets) but not with others (Carlson & de Mamani, 2010; Davidson et al., 2004; Strunk et al., 2010). For example, Davidson et al. (2004) assessed patient outcomes using self and observer ratings of depressive and anxiety symptoms, social functioning, global functioning, and number of episodes of deliberate self-harm. They reported that at 6-month follow up, there was a significant association between measured therapy quality and observer-rated depression only; at 12-month follow-up, there were significant associations between therapy quality and all observer-rated clinical outcomes but not for self-rated outcome measures or the number of self-harm episodes during follow-up. Interestingly, the one study that involved non-specialists in this setting reported no change in patient outcomes after evaluation of therapist competency (Butler et al., 2013). All studies conducted in LMIC settings reported positive associations between therapy quality and patient outcomes (Chowdhary et al., 2016; Nadkarni et al., 2017; Papas et al., 2011).

4. Discussion

Our objective was to conduct a scoping review of the association of therapist competency with client treatment outcomes, with an emphasis

on the use of standardized role plays. We identified four reviews encompassing a total of 61 studies on this topic. The reviews summarized moderate to no association between competency and treatment outcomes. Within these reviews, the majority of studies did not operationalize competency with observed standardized role plays as recommended by (Fairburn & Cooper (2011)). Instead, the studies used ratings of actual client sessions, which is consistent with Fairburn and Cooper's definition of therapy quality, i.e., how care is delivered under real world circumstances. Therefore, the reviews do not allow us to draw a conclusion regarding the relationship between competency—as measured by standardized role plays—with client outcomes.

For our second objective, we turned to literature on competency tool development to determine if these studies could elucidate the relationship between competency and client outcomes. We identified 39 tools reportedly designed to evaluate competency. These 39 tools had been used in 43 studies addressing competency, and 21 of the studies reported a comparison between competency and client outcomes. However, as with the systematic reviews, we found that few studies evaluated competency in a manner consistent with Fairburn and Cooper's recommendation to use structured role plays. Of the 43 studies, only 8 described using the tool to rate structured role plays, and none of these studies presented findings comparing performance on these role plays and client outcomes.

Of the 21 studies that did report an association between competency and client outcomes, all of them used the tool to rate actual sessions, i.e., generating therapy quality ratings. Most studies were conducted in high-income countries and assessed specialist therapists. These studies yielded a mixed picture with 8 studies demonstrating a positive association between therapy quality and client outcomes, 5 reporting mixed results, and 4 with no effect (see Fig. 4). Interestingly, one study
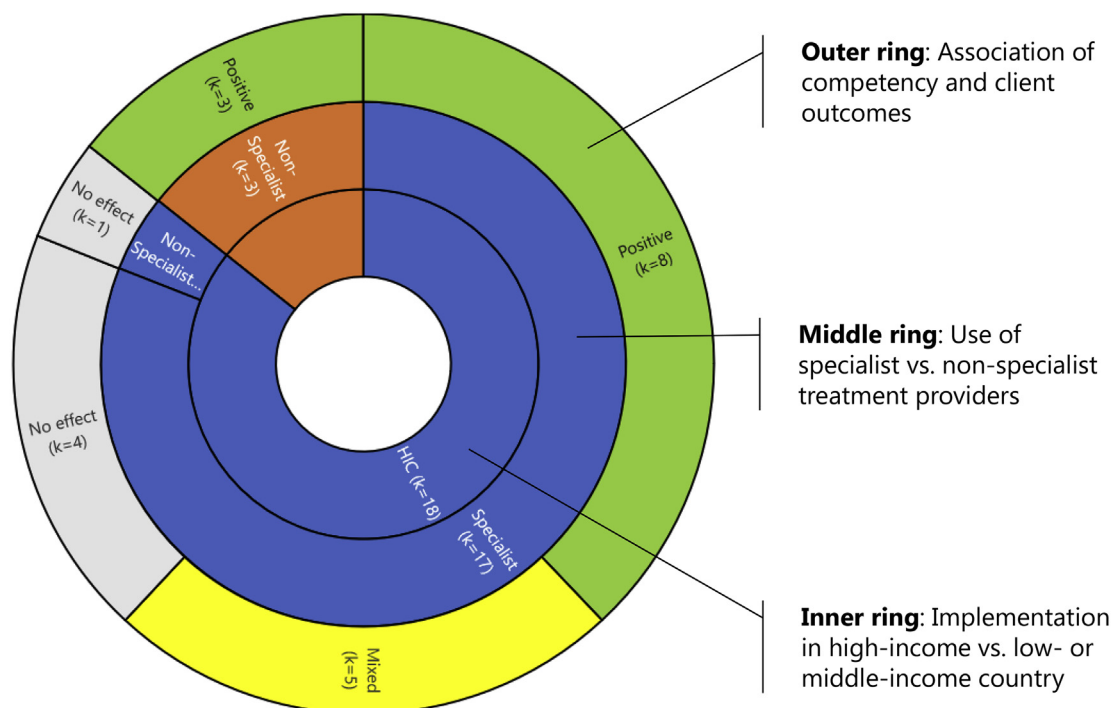
Fig. 4. Chart of Findings for Evaluated Therapist Competency and Client Outcomes.
HIC, high-income countries; LMIC, low- and middle-income countries.

conducted with non-specialists in a HIC setting reported no effect. All three of the LMIC studies reported a positive association, and notably these were all with non-specialists.

Rating in these studies was typically done through use of an independent rater system to assess a randomly-chosen selection of video-recorded or audio-recorded client sessions. While the evaluation of therapy quality is a proxy for therapist competency, one of the challenges in extrapolating competency from actual client sessions is that few therapy sessions (or segments of sessions) may make it difficult to conclude if a therapist has achieved full competency in all skills needed for appropriate care. This is especially relevant from a training perspective. Three reviews highlighted that such use of randomly selected sessions may lose the potentially important causal relations from contextual background or processes that occur in the intervening periods between sessions (Collyer et al., 2019; Kazantzis, 2003; Webb et al., 2010). Another finding was that more than half of the rating scales used Likert response with 7 or more levels. This creates a high cognitive burden that may limit feasibility of use by non-specialists scoring the rating tools. In addition, the high number of response options may create challenges to achieving adequate inter-rater reliability (Kohrt, Ramaiya, Rai, Bhardwaj, & Jordans, 2015b).

The large amount of heterogeneity in the field is due to the diversity in the methods of modeling used, range of assessment tools, sample sizes of therapists, different outcome measures, and evaluation methods (live, audio, video, etc.). Most studies opted to measure a multitude of outcomes at once (Butler et al., 2013; Carlson & de Mamani, 2010; Gutermann et al., 2015; Hogue et al., 2008b; Martino et al., 2008; Paivio et al., 2004; Papas et al., 2011) while others chose only one or two outcomes, but opted for several different measurement methods for the same outcome (Davidson et al., 2004; Hoffart et al., 2005; Strunk et al., 2010; Trepka et al., 2004). Furthermore, much of the literature that assesses the relation between therapist competency and client outcomes has been derived from specialist therapists in HIC. The lack of a relationship between competency and outcomes therefore may be due to a lack of variation in competency in these relatively well-trained samples.

In addition to low therapist sample sizes, there were also small client sample sizes, with some including as few as 8 clients (Dittman et al., 2017) but others ranging to almost 500 (Patel et al., 2017). With such low numbers of participants in certain studies, the statistical power of these studies may be too small to capture a significant effect on client outcomes. Although the sample sizes in the HIC studies did not seem to drive outcomes in either direction, the LMIC studies reported positive results with comparatively large samples of 75 participants (Papas et al., 2011), 377 (Nadkarni et al., 2017) and 495 (Patel et al., 2017), respectively.

Another potential drawback of most study designs is that all reviews statistically modeled their results with the assumption of a linear correlation between therapist competency and client outcomes. The relationship with competency may instead follow a non-linear relationship that is not captured with analytic techniques such as correlation and linear regression. Competency below a certain threshold may be associated with lack of client benefit, but competency above that threshold may not show additional gains regarding client benefits. Based on this type of relationship, if most therapists engaged in these studies had, in principle, achieved minimum competency standards then there may have been limited range for statistical comparisons (i.e., lacking therapists below the threshold).

Finally, the highest quality of delivery of certain interventions may not be what is related to expected outcomes; some more moderate levels of competency could be more predictive of improvement. In HICs, competency in certain common factors can contribute to this "primary care paradox," or the observation that generalists can treat some conditions well despite delivering manualized care with less technical proficiency than a specialist (Stange & Ferrer, 2009); i.e., highlighting the importance of simultaneously understanding "the value and quality of care at the level of specific illnesses, whole people, communities, and populations," as these "different levels may have different drivers of process and outcome." Although current literature is sparse, the implication that non-specialists can deliver interventions "well enough" and produce positive outcomes is promising for treating those populations that otherwise may not have as much access to psychological care. Determining which demonstrated provider attributes are most effective for influencing client change for various disorders may therefore be of

interest for further investigation.

Several competency assessment tools were heavily focused on common factors (Allison, Bes, & Rose, 2012; Decker, Nich, Carroll, & Martino, 2014; Miller, Moyers, Ernst, & Amrhein, 2008; Moyers, Martin, Manuel, Miller, & Ernst, 2010), but the majority of the tools focused on measuring specialist techniques. However, every competency assessment tool in our review included at least one common factor item. Global items such as empathy, collaboration, and interpersonal effectiveness were mentioned as strong therapist factors on positive client outcomes in several studies (Gaume et al., 2008; Ginzburg et al., 2012; Spohr et al., 2016). In a study using the Manual for Motivational Interviewing Skill Code (v 2.1), the authors reported there was no statistically significant relationship between therapist quality and overall client outcomes; however, they noted that a counselor's empathy was significantly correlated with a client's weekly alcohol use decrease and a decreased number of heavy drinking episodes (Gaume et al., 2008). Another study indicated that despite therapist quality not having a statistically significant effect on client substance use at the end of the study, therapists' demonstration of empathy was highly indicative of clients' early treatment efforts (Spohr et al., 2016).

Although most studies only described outcomes in relation to their overall scores on the competency assessment tool, some indicated treatment-specific techniques that were proven to have a higher influence over outcomes than other factors. Items varied by therapy type; Motivational Interviewing (MI) studies emphasized "evocation" and "autonomy support" (Gaume et al., 2008; Spohr et al., 2016); an adapted Cognitive Behavioral Therapy for Social Phobia Scale (CTCS-SP) mentioned factors such as "self-focused attention and imagery" or "resource activation" as predictors of positive patient outcomes (Ginzburg et al., 2012). One study using the Cognitive Therapy Scale-Revised (CTS-R) indicated that although the overall therapist quality scores resulted in a weak, positive association with improved client outcomes, the "specific techniques" subscale had the highest correlation with these outcomes.

One of the limitations of our scoping review, following the suggestions of Arksey and O'Malley (2005), was that we did not assess the quality of studies included.

4.1. Competency Research agenda and future directions for the field

As our findings demonstrate, there is a need for consistently-defined, reliable methods to assess therapist competency. We propose a research agenda for the competency field (see Textbox 1). The key issues will be to use feasible approaches consistent with expert recommendations regarding standardized, role play-based strategies. In addition, analytic and design techniques should consider that competency and client outcomes may not always be associated in a linear relationship; instead, a threshold-based relationship may be involved. Therefore, studies will need to be designed in a manner that includes providers below the competency threshold. This poses ethical challenges and would lend itself to using health systems data from providers rather than setting up experimental conditions with potentially incompetent providers. Within health systems, providers could be evaluated using competency assessments with standardized role plays, and this could then be compared with clinical outcomes of their client populations. Fig. 5 outlines potential conceptual models for designing

competency studies, including moderation and mediation relationships. Competency in these models could be considered in categorical (threshold) based approaches in addition to the linear associations that have previously been tested.

A major critique of this recommended use of structured role plays with simulated clients is the question of the time and resource allocation necessary to implement this methodology. Although role play methods may be currently more resource-intensive, the alternative has a host of limitations. Considering that the current research has been primarily limited to high-resource settings, applied knowledge testing has been a proposed cost-effective alternative to measurement of provider competency. However, reviews of these applied tests thus far have suggested that the measurements employed tend to be narrowly focused on their specific treatments' strategies and procedures (Herschell, Kolko, Baumann, & Davis, 2010; Rakovshik & McManus, 2010). Their use in populations with limited literacy may also result in the loss of intended meaning of psychological concepts. Therefore, observing human behavior may be more appropriate. Part of the overarching goal of understanding competency, especially with non-specialist providers, is to determine a range of strategies for role plays and low-resource approaches to assessing behavior. Further research is needed to generate cost-effective ways to utilize role plays.

Ultimately, developing feasible and valid ways to assess competency is a crucial step toward developing provider evaluation and certification processes that provide assurance of safe and effective care. This review was part a World Health Organization (WHO) initiative: *Ensuring Quality in Psychological* Support *(EQUIP)* platform (https://www.who.int/mental_health/emergencies/equip/en/). The goal of EQUIP is to facilitate competency-based training in psychosocial support and psychological treatments (Kohrt et al., 2020). The EQUIP platform aligns with WHO's work on Universal Health Coverage that is establishing competency frameworks across fields of healthcare. EQUIP will be an online resource to help program managers and trainers utilize competency assessments to evaluate trainings and to feedback those competency results to support trainee development and modify curricula. EQUIP will encompass a competency evaluation tool, the Enhancing Assessment of Common Therapeutic Factors (ENACT), that has been developed for role play based assessment of mental health and psychosocial support skills for non-specialist and specialist providers across cultures, context, and types of interventions (Kohrt et al., 2015a, 2018). The full EQUIP platform will include information for implementing competency assessments including how to achieve inter-rater reliability with global rating standards and how to use role plays to assess competency.

5. Conclusion

While there is growing interest in identifying and measuring therapist competency in psychological treatments, varied approaches are represented in the current literature. Common sense dictates that a competent therapist would produce better client outcomes, but the heterogeneity of both therapist assessment methods and evaluations of client outcomes poses a challenge to comparing the results of these studies effectively. The absence of distinctions between the measurement of therapist *competency* and therapist *quality* muddles the ability for accurate cross-comparisons among studies. The lack of adequate

Textbox 1
Recommendations for a competency research agenda.

---

1. Competency should be defined consistently across studies, e.g., with Fairburn & Cooper (2011) as a standard definition.
2. Competency should be assessed with a consistent methodology, e.g., standardized role plays.
3. Competency evaluations done through standardized role plays should be distinct from quality and fidelity as assessed in actual sessions
4. Competency comparisons with client outcomes should employ designs that adequately address potential non-linear relationships (e.g., thresholds) and variability (range) of competency for providers in the study

---

a. **Prospective correlational analysis**, e.g., greater therapist competency predicts better client outcomes
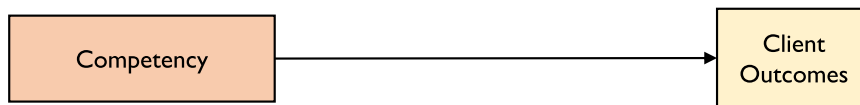
b. **Moderation analysis**, e.g., differences in treatment conditions (such as experimental vs. control treatments) are observed only when therapists have minimally adequate competency
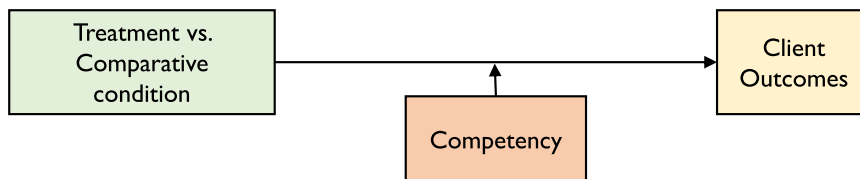
c. **Mediation analysis**, e.g., differences in competency explain a portion of the variance between training/supervision strategies and client outcomes
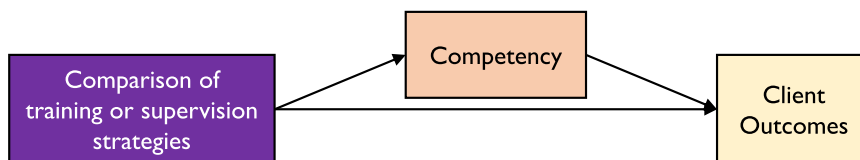
d. **Mediated mediation analysis**, e.g., a portion of the variance between competency and client outcomes is explained by differences in quality, with greater competency associated with greater quality and ultimately better client outcomes
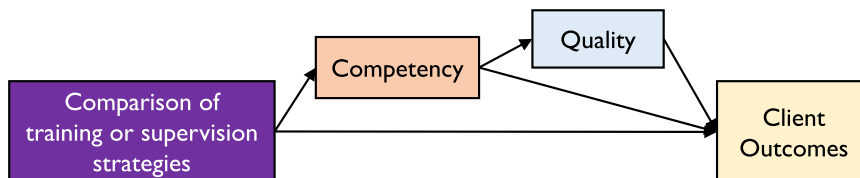
Fig. 5. Conceptual models for designing competency evaluation studies.

evaluation of therapist competency further hinders the ability to determine the skill level required to identify a provider as sufficiently "competent" in treatment delivery. The need for reliable, standardized methods to assess therapist *competency* prior to treating clients remains a significant gap in the literature. This is becoming increasingly important as more focus is being paid to involving non-specialists in psychological care, especially in low-resource settings. The results of the several studies reviewed here that evaluate non-specialist therapist quality in LMICs are promising for future research. Going forward, a new research agenda is needed for measuring competency in a manner consistent with expert recommendations that will be feasible and valid for assuring safe and effective care for populations around the world.

Declaration of competing interest

The authors declare no conflicts of interest.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.brat.2019.103531.

References

Allison, J., Bes, R., & Rose, G. (2012). Motivational Interviewing Target Scheme | MITS 2 . 1. *Exp. Guid. Contents*, 1–26 (July).

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology: Theory and Practice, 8*(1), 19–32. https://doi.org/10.1080/1364557032000119616.

Barber, J. P., Sharpless, B. A., Klostermann, S., & McCarthy, K. S. (2007). Assessing intervention competence and its relation to therapy outcome: A selected review derived from the outcome literature. *Professional Psychology: Research and Practice, 38*(5),

493.

Barnett, M. L., Lau, A. S., & Miranda, J. (2018). Lay health worker involvement in evidence-based treatment delivery: A conceptual model to address disparities in care. *Annual Review of Clinical Psychology, 14*, 185–208. https://doi.org/10.1146/annurev-clinpsy-050817-084825.

Bryant, R. A., Schafer, A., Dawson, K. S., Anjuri, D., Mulili, C., Ndogoni, L., et al. (2017). Effectiveness of a brief behavioural intervention on psychological distress among women with a history of gender-based violence in urban Kenya: A randomised clinical trial. *PLoS Medicine, 14*(8), e1002371. https://doi.org/10.1371/journal.pmed.1002371.

Butler, C. C., Simpson, S. A., Hood, K., Cohen, D., Pickles, T., Spanou, C., et al. (2013). Training practitioners to deliver opportunistic multiple behaviour change counselling in primary care: A cluster randomised trial. *BMJ, 346*. https://doi.org/10.1136/bmj.f1191 f1191-f1191.

Campos-Melady, M., Smith, J. E., Meyers, R. J., Godley, S. H., & Godley, M. D. (2017). The effect of therapists' adherence and competence in delivering the adolescent community reinforcement approach on client outcomes. *Psychology of Addictive Behaviors, 31*(1), 117–129. https://doi.org/10.1037/adb0000216.

Carlson, R. G., & de Mamani, A. G. W. (2010). Client characteristics and therapist competence and adherence to family therapy for schizophrenia. *Revista Interamericana de Psicología, 44*(2), 342–351.

Chowdhary, N., Anand, A., Dimidjian, S., Shinde, S., Weobong, B., Balaji, M., et al. (2016). The Healthy activity program lay counsellor delivered treatment for severe depression in India: Systematic development and randomised evaluation. *British Journal of Psychiatry, 208*(4), 381–388. https://doi.org/10.1192/bjp.bp.114.161075.

Collyer, H., Eisler, I., & Woolgar, M. (2019). Systematic literature review and meta-analysis of the relationship between adherence, competence and outcome in psychotherapy for children and adolescents. *European Child & Adolescent Psychiatry,* 1–15. https://doi.org/10.1007/s00787-018-1265-2.

Cuijpers, P., Reijnders, M., & Huibers, M. J. H. (2019). *The role of common factors in psychotherapy outcomes.* 1–25. https://doi.org/10.1146/annurev-clinpsy-050718-095424.

Davidson, K., Scott, J., Schmidt, U., Tata, P., Thornton, S., & Tyrer, P. (2004). Therapist competence and clinical outcome in the prevention of parasuicide by manual assisted cognitive behaviour therapy trial: The POPMACT study. *Psychological Medicine, 34*(5), 855–863.

Decker, S. E., Nich, C., Carroll, K. M., & Martino, S. (2014). Development of the therapist empathy scale. *Behavioural and Cognitive Psychotherapy, 42*, 339–354. https://doi.org/10.1017/S1352465813000039 03.

Dittman, C., Muller-Engelmann, M., Stangier, U., Priebe, K., Fydrich, T., Gorg, N., et al. (2017). Disorder- and treatment-specific therapeutic competence scales for post-traumatic stress disorder intervention: Development and psychometric properties. *Journal of Traumatic Stress, 30*, 614–625. https://doi.org/10.1002/jts.

Fairburn, C., & Cooper, Z. (2011). Therapist competence, therapy quality, and therapist training. *Behaviour Research and Therapy, 49*(6–7), 373–378. https://doi.org/10.1016/j.brat.2011.03.005.

Fairburn, C. G., & Patel, V. (2014). The global dissemination of psychological treatments: A road map for research and practice. *American Journal of Psychiatry, 171*(5), 495–498.

Forgatch, M. S., Patterson, G. R., & DeGarmo, D. S. (2005). Evaluating fidelity: Predictive validity for a measure of competent adherence to the Oregon model of parent management training. *Behavior Therapy, 36*(1), 3–13.

Fuhr, D. C., Weobong, B., Lazarus, A., Vanobberghen, F., Weiss, H. A., Singla, D. R., et al. (2019). Delivering the thinking Healthy Programme for perinatal depression through peers: An individually randomised controlled trial in India. *The Lancet Psychiatry, 6*(2), 115–127. https://doi.org/10.1016/S2215-0366(18)30466-8.

Gaume, J., Gmel, G., & Daeppen, J. B. (2008). Brief alcohol interventions: Do counsellors' and patients' communication characteristics predict change? *Alcohol and Alcoholism, 43*(1), 62–69. https://doi.org/10.1093/alcalc/agm141.

Gaume, J., Gmel, G., Faouzi, M., & Daeppen, J.-B. (2009). Counselor skill influences outcomes of brief motivational interventions. *Journal of Substance Abuse Treatment, 37*(2), 151–159. https://doi.org/10.1016/j.jsat.2008.12.001.

Global Social Service Workforce Alliance Interest Group on Para Professionals. (2017). *Para professionals in the social service workforce: Guiding principles, functions and competencies* (2nd). Global Social Service Workforce Alliancehttp://www.socialserviceworkforce.org/resources/para-professionals-social-service-workforce-guiding-principles-functions-and-0.

van Ginneken, N., Tharyan, P., Lewin, S., Rao, G. N., Meera, S., Pian, J., et al. (2013). Non-specialist health worker interventions for the care of mental, neurological and substance-abuse disorders in low-and middle-income countries. *Cochrane Database of Systematic Reviews, 11*, CD009149.

Ginzburg, D. M., Bohn, C., Hofling, V., Weck, F., Clark, D. M., & Stangier, U. (2012). Treatment specific competence predicts outcome in cognitive therapy for social anxiety disorder. *Behaviour Research and Therapy, 50*(12), 747–752. https://doi.org/10.1016/j.brat.2012.09.001.

Gutermann, J., Schreiber, F., Matulis, S., Stangier, U., Rosner, R., & Steil, R. (2015). Therapeutic adherence and competence scales for developmentally adapted cognitive processing therapy for adolescents with PTSD. *European Journal of Psychotraumatology, 6*(1), https://doi.org/10.3402/ejpt.v6.26632.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando, Florida: Academic Press.

Herschell, A. D., Kolko, D. J., Baumann, B. L., & Davis, A. C. (2010). The role of therapist training in the implementation of psychosocial treatments: A review and critique with recommendations. *Clinical Psychology Review, 30*(4), 448–466. https://doi.org/10.1016/j.cpr.2010.02.005.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ, 327*(7414), 557. https://doi.org/10.1136/bmj.327.7414.557.

Hoffart, A., Sexton, H., Nordahl, H. M., & Stiles, T. C. (2005). Connection between patient and therapist and therapist's competence in schema-focused therapy of personality problems. *Psychotherapy Research, 15*(4), 409–419. https://doi.org/10.1080/10503300500091702.

Hogue, A., Dauber, S., Chinchilla, P., Fried, A., Henderson, C. E., Inclan, J., et al. (2008a). Assessing fidelity in individual and family therapy for adolescent substance abuse. *Journal of Substance Abuse Treatment, 35*(2), 137–147. https://doi.org/10.1016/j.jsat.2007.09.002.

Hogue, A., Henderson, C. E., Dauber, S., Barajas, P. C., Fried, A., & Liddle, H. A. (2008b). Treatment adherence, competence, and outcome in individual and family therapy for adolescent behavior problems. *Journal of Consulting and Clinical Psychology, 76*, 544–555. https://doi.org/10.1037/0022-006X.76.4.544.

Kakuma, R., Minas, H., van Ginneken, N., Dal Poz, M. R., Desiraju, K., Morris, J. E., et al. (2011). Human resources for mental health care: Current situation and strategies for action. *The Lancet, 378*(9803), 1654–1663.

Kazantzis, N. (2003). Therapist competence in cognitive-behavioural therapies: Review of the contemporary empirical evidence. *Behaviour Change, 20*, 1–12. https://doi.org/10.1375/bech.20.1.1.24845 01.

Kohrt, B. A., Jordans, M. J. D., Rai, S., Shrestha, P., Luitel, N. P., Ramaiya, M., et al. (2015a). Therapist competence in global mental health: Development of the enhancing assessment of common therapeutic factors (ENACT) rating scale. *Behaviour Research and Therapy, 69*, 11–21. https://doi.org/10.1016/j.brat.2015.03.009.

Kohrt, B. A., Mutamba, B. B., Luitel, N. P., Gwaikolo, W., Onyango Mangen, P., Nakku, J., et al. (2018). How competent are non-specialists trained to integrate mental health services in primary care? Global health perspectives from Uganda, Liberia, and Nepal. *International Review of Psychiatry, 30*(6), 182–198. https://doi.org/10.1080/09540261.2019.1566116.

Kohrt, B. A., Ramaiya, M. K., Rai, S., Bhardwaj, A., & Jordans, M. J. D. (2015b). Development of a scoring system for non-specialist ratings of clinical competence in global mental health: A qualitative process evaluation of the enhancing assessment of common therapeutic factors (ENACT) scale. *Global Mental Health, 2*, e23. https://doi.org/10.1017/gmh.2015.21.

Kohrt, B. A., Schafer, A., Willhoite, A., van 't Hof, E., Pedersen, G. A., Watts, S., et al. (2020). Ensuring quality in psychological support (WHO EQUIP): Developing a competent global workforce. *World Psychiatry, 19*(1), https://doi.org/10.1002/wps.20704 doi:DOI:10.1002/wps.20704.

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science, 5*(1), https://doi.org/10.1186/1748-5908-5-69 69-69.

Martino, S., Ball, S. A., Nich, C., Frankforter, T. L., & Carroll, K. M. (2008). Community program therapist adherence and competence in motivational enhancement therapy. *Drug and Alcohol Dependence, 96*(1), 37–48.

Mays, N., Roberts, E., & Popay, J. (2001). Synthesising research evidence. In N. Fulop, P. Allen, A. Clarke, N. Black, & Fulop (Eds.). *Methods for studying the delivery and organisation of health services*London: Routledge 194-194.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine, 65*(9), S63–S67.

Miller, W. R., Moyers, T. B., Ernst, D., & Amrhein, P. (2008). *Manual for the motivational interviewing skill Code (MISC), version 2.1. Center on alcoholism, substance abuse, and addictions.* The University of New Mexico.

Möller, C., Karlgren, L., Sandell, A., Falkenström, F., & Philips, B. (2017). Mentalization-based therapy adherence and competence stimulates in-session mentalization in psychotherapy for borderline personality disorder with co-morbid substance dependence. *Psychotherapy Research, 27*(6), 749–765. https://doi.org/10.1080/10503307.2016.1158433.

Montgomery, E. C., Kunik, M. E., Wilson, N., Stanley, M. A., & Weiss, B. (2010). Can paraprofessionals deliver cognitive-behavioral therapy to treat anxiety and depressive symptoms? *Bulletin of the Menninger Clinic, 74*(1), 45–62.

Moyers, T. B., Martin, T., Manuel, J. K., Miller, W. R., & Ernst, D. (2010). *Revised global scales: Motivational interviewing treatment integrity 3.1.1 (MITI 3.1.1). Center on alcoholism, substance abuse, and addictions, Vol. 1,* The University of New Mexico1–29 January.

Mulder, R., Murray, G., & Rucklidge, J. (2017). Common versus specific factors in psychotherapy: Opening the black box. *The Lancet Psychiatry, 4*(12), 953–962. https://doi.org/10.1016/S2215-0366(17)30100-1.

Muse, K., & McManus, F. (2013). A systematic review of methods for assessing competence in cognitive–behavioural therapy. *Clinical Psychology Review, 33*(3), 484–499. https://doi.org/10.1016/j.cpr.2013.01.010.

Nadkarni, A., Weobong, B., Weiss, H. A., McCambridge, J., Bhat, B., Katti, B., et al. (2017). Counselling for alcohol problems (CAP), a lay counsellor-delivered brief psychological treatment for harmful drinking in men, in primary care in India: A randomised controlled trial. *Lancet, 389*(10065), 186–195. https://doi.org/10.1016/s0140-6736(16)31590-2.

National Collaborating Centre for Mental Health (2019). *The Improving Access to Psychological Therapies Manual.* London, United Kingdom: National Institute for Clinical Excellence, National Health Service. www.england.nhs.uk/publication/the-improving-access-to-psychological-therapies-manual/.

Paivio, S. C., Holowaty, K. A. M., & Hall, I. E. (2004). The influence of therapist adherence and competence on client reprocessing of child abuse memories. *Psychotherapy: Theory, Research, Practice, Training, 41*(1), 56–68. https://doi.org/10.1037/0033-3204.41.1.56.

Papas, R. K., Sidle, J. E., Gakinya, B. N., Baliddawa, J. B., Martino, S., Mwaniki, M. M., et al. (2011). Treatment outcomes of a stage 1 cognitive-behavioral trial to reduce alcohol use among human immunodeficiency virus-infected out-patients in western

Kenya. *Addiction, 106*(12), 2156–2166. https://doi.org/10.1111/j.1360-0443.2011. 03518.x.

Patel, V. (2012). Global mental health: From science to action. *Harvard Review of Psychiatry, 20*(1), 6–12.

Patel, V., Saxena, S., Lund, C., Thornicroft, G., Baingana, F., Bolton, P., et al. (2018). The Lancet Commission on global mental health and sustainable development. *The Lancet, 392*(10157), 1553–1598. https://doi.org/10.1016/S0140-6736(18)31612-X.

Patel, V., Weobong, B., Weiss, H. A., Anand, A., Bhat, B., Katti, B., et al. (2017). The Healthy activity program (HAP), a lay counsellor-delivered brief psychological treatment for severe depression, in primary care in India: A randomised controlled trial. *Lancet, 389*(10065), 176–185. https://doi.org/10.1016/s0140-6736(16) 31589-6.

Rahman, A., Hamdani, S., Awan, N., et al. (2016). Effect of a multicomponent behavioral intervention in adults impaired by psychological distress in a conflict-affected area of Pakistan: A randomized clinical trial. *Journal of the American Medical Association, 316*(24), 2609–2617. https://doi.org/10.1001/jama.2016.17165.

Rahman, A., Khan, M. N., Hamdani, S. U., Chiumento, A., Akhtar, P., Nazir, H., et al. (2019). Effectiveness of a brief group psychological intervention for women in a post-conflict setting in Pakistan: A single-blind, cluster, randomised controlled trial. *The Lancet, 393*(10182), 1733–1744.

Rakovshik, S. G., & McManus, F. (2010). Establishing evidence-based training in cognitive behavioral therapy: A review of current empirical findings and theoretical guidance. *Clinical Psychology Review, 30*(5), 496–516. https://doi.org/10.1016/j.cpr.2010.03. 004.

Resnick, B., Inguito, P., Orwig, D., Yahiro, J. Y., Hawkes, W., Werner, M., et al. (2005). *Treatment fidelity in behavior change research: A case example.* Haggerstown, MD: Lippincott Williams & Wilkins, Inc139–143.

Sharpless, B. A., & Barber, J. P. (2009). *A conceptual and empirical review of the meaning, measurement, development, and teaching of intervention competence in clinical psychology.*

Shaw, B. F. B., Elkin, I. I., Yamaguchi, J. J., Olmsted, M. M., Vallis, T. M. T., Dobson, K. S. K., et al. (1999). Therapist competence ratings in relation to clinical outcome in cognitive therapy of depression. *Journal of Consulting and Clinical Psychology, 67*(6), 837–846.

Sikander, S., Ahmad, I., Atif, N., Zaidi, A., Vanobberghen, F., Weiss, H. A., et al. (2019). Delivering the thinking Healthy Programme for perinatal depression through volunteer peers: A cluster randomised controlled trial in Pakistan. *The Lancet Psychiatry, 6*(2), 128–139. https://doi.org/10.1016/S2215-0366(18)30467-X.

Singla, D. R., Kohrt, B. A., Murray, L. K., Anand, A., Chorpita, B. F., & Patel, V. (2017). Psychological treatments for the world: Lessons from low- and middle-income countries. *Annual Review of Clinical Psychology, 13*(April), 5.1–5.33. https://doi.org/ 10.1146/annurev-clinpsy-032816-045217.

Spohr, S. A., Taxman, F. S., Rodriguez, M., & Walters, S. T. (2016). Motivational interviewing fidelity in a community corrections setting: Treatment initiation and subsequent drug use. *Journal of Substance Abuse Treatment, 65*, 20–25. https://doi.org/ 10.1016/j.jsat.2015.07.012.

Stange, K. C., & Ferrer, R. L. (2009). The paradox of primary care. *The Annals of Family Medicine, 7*(4), 293–299.

Strunk, D. R., Brotman, M. A., DeRubeis, R. J., & Hollon, S. D. (2010). Therapist competence in cognitive therapy for depression: Predicting subsequent symptom change. *Journal of Consulting and Clinical Psychology, 78*(3), 429–437. https://doi.org/10. 1037/a0019631.

Trepka, C., Rees, A., Shapiro, D. A., Hardy, G. E., & Barkham, M. (2004). Therapist competence and outcome of cognitive therapy for depression. *Cognitive Therapy and Research, 28*(2), 143–157. https://doi.org/10.1023/B:COTR.0000021536.39173.66.

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., et al. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine, 169*(7), https://doi.org/10.7326/M18-0850 467-467.

UCL Centre for Outcomes, Research and Effectiveness. (2019). *Assessing competences against the cognitive behaviour therapy framework self-assessment tool.* London, UK: University College London. https://www.ucl.ac.uk/CORE/, Accessed date: 2 February 2019.

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry, 14*(3), 270–277. https://doi.org/10.1002/wps.20238.

Webb, C. A., Derubeis, R. J., & Barber, J. P. (2010). Therapist adherence/competence and treatment outcome: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 78*(2), 200–211.